

# ON THE ESTIMATION OF THE CREDIBILITY FACTOR: A BAYESIAN APPROACH\*

BY RENÉ SCHNIEPER

*“Zürich” Versicherungs-Gesellschaft, Zürich*

## ABSTRACT

In practical applications of Credibility Theory the structure parameters usually have to be estimated from the data. This leads to an estimator of the a posteriori mean which is often biased and where the credibility factor depends on the data. A more coherent approach to the problem would be to also treat the unknown parameters as random variables and to simultaneously estimate the a posteriori mean and the structure parameters. Different statistical models are proposed which allow for such a solution. These models all lead to an estimation of the posterior mean which is a weighted average of the prior mean and of the observed mean, the weights depending on the observations.

## KEYWORDS

Credibility theory; Structure parameters; Statistical estimation; exponential families; Pricing of risks.

## 1. INTRODUCTION

We have  $k$  different risks with a claims record over a certain number of years

$$X_{11}, X_{12}, \dots, X_{1n_1}$$

$$X_{21}, X_{22}, \dots, X_{2n_2}$$

$$\vdots$$

$$X_{k1}, X_{k2}, \dots, X_{kn_k}$$

Depending on the specific problem, the data are numbers of claims from different insurance policies, loss ratios from insurance portfolios for instance in fire, liability or workmen's compensation insurance and burning costs from different reinsurance treaties.

With each claims record  $X_{ij}$  there is an associated measure of risk exposure  $p_{ij}$  which is a number of risk years, a sum insured, a turnover, a total amount of wages or a premium income depending on the specific problem.

---

\* Paper presented at the XXIVth ASTIN Colloquium in Cambridge.

Within the framework of credibility theory it is assumed that

1. Each risk  $i$  is characterized by an individual unknown risk parameter  $\theta_i$ ; the risk parameters  $\theta_1, \theta_2, \dots, \theta_k$  are i.i.d. random variables.  
The random vectors  $(\theta_i, X_{i1}, \dots, X_{in_i})$   $i = 1, \dots, k$  are independent.
2. Given  $\theta_i$  the observations

$$X_{i1}, X_{i2}, \dots, X_{in_i}$$

are independent with finite second moments

$$E(X_{ij} | \theta_i) = \mu(\theta_i)$$

$$\text{Var}(X_{ij} | \theta_i) = \frac{\sigma^2(\theta_i)}{p_{ij}}$$

The individual premium  $\mu(\theta_i)$  is to be approximated by a premium which is linear in the observations

$$\hat{\mu}(\theta_i) = \alpha_{i0} + \sum_{j=1}^{n_i} \alpha_{ij} X_{ij}$$

and which minimizes the expected squared error

$$E[\mu(\theta_i) - \hat{\mu}(\theta_i)]^2$$

It is shown that the optimal linear premium is a weighted average of the individual mean and of the a priori mean

$$\hat{\mu}(\theta_i) = z_i \bar{x}_i + (1 - z_i) m$$

where

$$m = E[\mu(\theta)]$$

$$\bar{x}_i = \left( \sum_{j=1}^{n_i} p_{ij} x_{ij} \right) / \left( \sum_{j=1}^{n_i} p_{ij} \right)$$

The weight  $z_i$  given to the individual mean  $\bar{x}_i$  is called the credibility factor. It is equal to

$$z_i = \frac{p_i \cdot b}{p_i \cdot b + w}$$

where

$$p_i = \sum_{j=1}^{n_i} p_{ij}$$

$$b = \text{Var}[\mu(\theta)]$$

$$w = E[\sigma^2(\theta)]$$

$b$  and  $w$  are the between risks and within risks variance components respectively. To practically compute the credibility premium we need to estimate the structure parameters  $m$ ,  $b$  and  $w$ .

$m$  is either known a priori e.g. from some nationwide statistics or is estimated from the collective. The best linear unbiased estimator for  $m$  is

$$\hat{m} = \sum_{i=1}^k \frac{z_i}{z} \bar{x}_i.$$

with

$$z = \sum_{i=1}^k z_i.$$

In what follows we shall assume that  $m$  is known a priori and focus on the estimation of  $b$  and  $w$ . [In KLUGMAN (1986)  $m$  is assumed to be a random variable and included in the Bayesian analysis].

In their pioneering paper BÜHLMAN and STRAUB (1970) propose the following “natural” estimators.

$$\begin{aligned}\hat{w} &= \frac{1}{k} \sum_{i=1}^k \frac{1}{n_i - 1} \sum_j p_{ij} (x_{ij} - \bar{x}_i)^2 \\ \hat{b} &= \frac{1}{c} \left( \sum_{i=1}^k \frac{p_i}{p} (\bar{x}_i - \bar{x})^2 - (k-1) \frac{\hat{w}}{p} \right)\end{aligned}$$

where

$$\begin{aligned}p &= \sum_i p_i \\ \bar{x} &= \left( \sum_i p_i \bar{x}_i \right) / p \\ c &= \sum_i \frac{p_i}{p} \left( 1 - \frac{p_i}{p} \right)\end{aligned}$$

Since the estimator of the between risks variance component may be negative, it is replaced by  $\max(\hat{b}, 0)$  in practical applications.

Other estimators have also been proposed; for a review of the subject see for instance DUBEY and GISLER (1981). A common property of these estimators is that they usually lead to a biased estimator of  $\mu(\theta)$ .

We shall adopt a different approach. Since  $b$  and  $w$  are unknown quantities we shall treat them as random variables and we shall propose statistical models which allow for a simultaneous estimation of  $b$ ,  $w$  and  $\mu(\theta)$ . The credibility factor will depend upon the observations, but this is also the case for the Bühlmann-Straub estimator once  $b$  and  $w$  have been replaced by  $\hat{b}$  and  $\hat{w}$ .

A general discussion of Bayesian inference on variance components in a normal model as well as in a generalisation of the normal model can be found in BOX and TIAO (1973).

The present paper focuses on applications to credibility theory.

## 2. THE BASIC MODEL

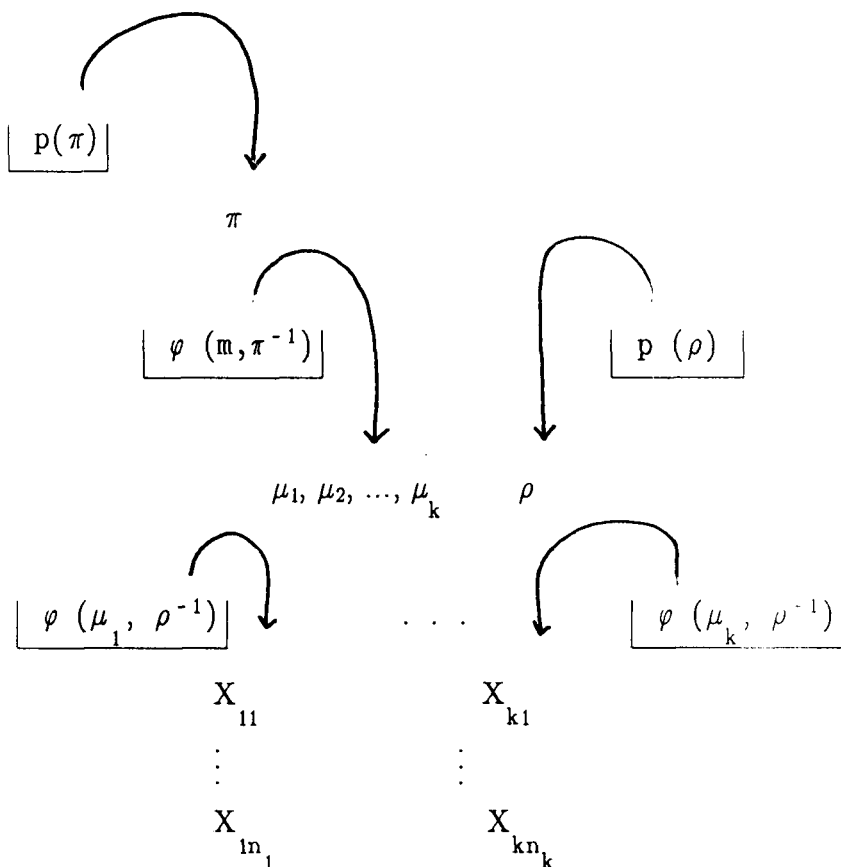
We assume that the data described at the outset is generated by the following model:

- the individual premiums  $\mu_i$  stem from a normal distribution with unknown precision  $\pi$ ;
- the claims record of risk  $i$

$$X_{i1}, X_{i2}, \dots, X_{in_i}$$

stems from a normal distribution with mean  $\mu_i$  and unknown precision  $\rho$ .

Note that instead of emphasizing the variance of a random variable, we emphasize its precision. This model is best illustrated by the following picture where the realizations of random variables are represented by drawings from an urn.



### Conventions and notation

- Densities are indexed by their arguments; thus  $p(\pi)$  and  $p(\rho)$  do not necessarily represent the same density.
- Distribution laws of random variables are symbolized by  $p(\cdot)$ . However we do not necessarily assume that the density of a random variable exists.
- $\varphi(m, \sigma^2)$  denotes the normal density with mean  $m$  and variance  $\sigma^2$ .

The formal model assumptions are as follows

- 1) The precision  $\pi$  of the individual premiums and the precision  $\rho$  of the observations are random variables whose distribution function will be specified later.
- 2) Each risk is characterized by an individual premium  $\mu_i$  and a common precision  $\rho$ .
  - i)  $\mu_1, \mu_2, \dots, \mu_k$  are independent random variables given  $\pi$ ; their common density is  $\varphi(m, \pi^{-1})$ .
  - ii)  $(\mu_i, x_{i1}, \dots, x_{in_i})$  are independent random vectors given  $\pi$  and  $\rho$ .
  - iii)  $\mu_1, \mu_2, \dots, \mu_k$  are independent random variables given  $\pi$ .
- 3) Given  $(\mu_i, \rho)$ ,  $x_{ij}$  ( $j = 1, \dots, n_i$ ) are independent random variables with common density  $\varphi(\mu_i, \rho^{-1})$ .

### Remarks

- 1) If the distribution of  $\pi$  and  $\rho$  were degenerated [i.e. if the probability mass of the common distribution of  $\pi$  and  $\rho$  is concentrated in some possibly unknown points  $(\pi_0, \rho_0)$ ] the above model would be a special case of the credibility model.
- 2) A similar model, but with individual precisions  $\rho_{ij}$  for each observation, has been proposed by JEWELL and SCHNIEPER (1983) for the treatment of outliers.

We now turn to the problem of the estimation of the individual premium. Let  $D$  denote the set of the claims records from all individual risks; the best estimator of the individual premium (best in the sense that it minimizes the expected squared error) is the posterior mean.

$$E(\mu_i | D) = \int E(\mu_i | D, \rho, \pi) p(\rho, \pi | D) d\rho \cdot d\pi$$

Because of assumptions 2) and 3) we have

$$E(\mu_i | D, \rho, \pi) = \frac{\pi m + \rho n_i \bar{x}_i}{\pi + \rho n_i} = x_i^*$$

because the credibility formula is exact in the case of a normal likelihood (with known variance) and a normal prior.

$\bar{x}_i$  is as in section 1 but with all measures of exposure equal to one.

Therefore we have

$$E(\mu_i | D) = E\left(\frac{\rho n_i}{\rho n_i + \pi} \mid D\right) \cdot \bar{x}_i + E\left(\frac{\pi}{\rho n_i + \pi} \mid D\right) \cdot m$$

$$E(\mu_i | D) = z_i(D) \bar{x}_i + [1 - z_i(D)] \cdot m$$

The posterior mean is given by a credibility type formula where the credibility factor depends on the data. To determine the credibility factor and the posterior mean, we must determine the posterior distribution of  $\pi$  and  $\rho$  given the data.

Using Bayes' Theorem we have

$$p(\rho, \pi | D) \propto p(D | \rho, \pi) p(\rho, \pi)$$

The common density of  $\rho$  and  $\pi$  will be specified later; for the likelihood we obtain

$$p(D | \rho, \pi) = \int p(D | \underline{\mu}, \rho, \pi) p(\underline{\mu} | \rho, \pi) d\mu_1 \dots d\mu_k$$

where

$$\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_k).$$

From assumption 2) ii) and 3) it follows

$$p(D | \underline{\mu}, \rho, \pi) \propto \rho^{\frac{1}{2} \sum_{i=1}^k n_i} \cdot e^{-\frac{1}{2} \rho \sum_j \sum_i (x_{ij} - \mu_i)^2}$$

and from 2) i)

$$p(\underline{\mu} | \rho, \pi) \propto \pi^{\frac{k}{2}} \cdot e^{-\frac{1}{2} \pi \sum_i (\mu_i - m)^2}$$

therefore the likelihood becomes

$$p(D | \rho, \pi) \propto \pi^{\frac{k}{2}} \cdot \rho^{\frac{1}{2} \sum_i n_i} \cdot e^{-\frac{1}{2} (\pi \sum_i (\mu_i - m)^2 + \rho \sum_j \sum_i (x_{ij} - \mu_i)^2)} d\mu_1 \dots d\mu_k$$

and upon integration we obtain

$$p(D | \rho, \pi) \propto \frac{\pi^{\frac{k}{2}} \rho^{\frac{1}{2} \sum_i n_i}}{\prod_i (\pi + \rho n_i)^{\frac{1}{2}}} \cdot e^{-\frac{1}{2} (\rho \sum_j (x_{ij}^2 - (\pi + \rho n_i) x_{ij}^* + \pi m^2))}$$

from which it is seen that

$$\sum_j x_{ij} \text{ and } \sum_j x_{ij}^2 \quad i = 1, \dots, k$$

are sufficient statistics.

After some straightforward but tedious algebraic transformations the likelihood can be written in an intuitively more appealing form

$$p(D|\rho, \pi) \propto \frac{\pi^{\frac{k}{2}} \rho^{\frac{1}{2} \sum_i n_i}}{\prod_i (\pi + \rho n_i)^{\frac{1}{2}}} \cdot e^{-\frac{1}{2} (\pi \sum_i (x_i^* - m)^2 + \pi \sum (x_{ij} - x_i^*)^2)}$$

from which we can compute the posterior density of  $\rho$  and  $\pi$  (once the prior density has been specified) and the posterior mean. The reason why the problem remains tractable is because  $p(D|\rho, \pi)$  is in analytical form; this in turn is due to the fact that for given  $\pi$  and  $\rho$   $p(D|\mu, \rho, \pi)$  and  $p(\mu|\rho, \pi)$  are conjugate priors.

### Remark

The following relation is true in general

$$E[\mu_i|D] = \int E[\mu_i|D, \rho, \pi] p(\rho, \pi|D) d\rho, d\pi,$$

and from credibility theory we obtain

$$E[\mu_i|D, \rho, \pi] \simeq \frac{\pi m + \rho n_i \bar{x}_i}{\pi + \rho n_i}$$

independently of the above distributional assumptions.

Therefore the form of the optimal estimator

$$E[\mu_i|D] \simeq z(D) \bar{x}_i + (1 - z(D)) m$$

is independent of the distributional assumptions; these are only needed to compute the credibility factor

$$z(D) = \int \frac{\rho n_i}{\pi + \rho n_i} p(\rho, \pi|D) d\rho d\pi.$$

### 3. A NUMERICAL ILLUSTRATION

We illustrate the results of section 2 with the following very simple numerical example

$$m = 0, \quad \pi = \frac{1}{3}$$

$$\rho = \begin{cases} 1 & \text{with probability 0.333} \\ 0.001 & \text{with probability 0.667} \end{cases}$$

$$E(\rho) = 0.334$$

There is only one risk and let us also assume that there is only one observation  $x$ .

$$\rho | D = \begin{cases} 1 & \text{with probability } p(D | \rho = 1) \cdot 0.333 \\ 0.001 & \text{with probability } p(D | \rho = 0.001) \cdot 0.667 \end{cases}$$

with

$$p(D | \rho) \propto \left( \frac{\rho}{0.333 + \rho} \right)^{\frac{1}{2}} e^{-\frac{1}{2} (0.333 x^*(\rho)^2 + \rho [x - x^*(\rho)]^2)}$$

$$x^*(\rho) = \frac{\rho x}{0.333 + \rho} = \begin{cases} 0.750 x & \text{for } \rho = 1 \\ 0.003 x & \text{for } \rho = 0.001 \end{cases}$$

As  $x$  becomes “large”, i.e. deviates strongly from  $m$  the whole probability mass is shifted towards  $\rho = 0.001$  and  $E(\mu | D)$  converges towards  $0.003x$ .

| $x$ | $P(\rho = 0.001   D)$ | $z(D)$ | $E(\mu   D)$ |
|-----|-----------------------|--------|--------------|
| 0   | 0.112                 | 0.666  | 0            |
| 1   | 0.125                 | 0.657  | 0.657        |
| 2   | 0.172                 | 0.621  | 1.243        |
| 3   | 0.279                 | 0.541  | 1.624        |
| 4   | 0.481                 | 0.391  | 1.564        |
| 5   | 0.739                 | 0.198  | 0.988        |
| 6   | 0.918                 | 0.064  | 0.387        |
| 7   | 0.983                 | 0.016  | 0.112        |
| 8   | 0.997                 | 0.005  | 0.040        |
| 9   | 1.000                 | 0.003  | 0.029        |
| 10  | 1.000                 | 0.003  | 0.030        |

where

$$z(D) = \left( \frac{\rho}{0.333 + \rho} \mid D \right)$$

$$E(\mu | D) = z(D) \cdot x + [1 - z(D)] \cdot m = z(D) \cdot x$$

If we take two observations instead of one a more interesting picture emerges. Two “large” observations are given a high credibility factor because the within variance component is low, whereas one “small” and one “large” observation are given a low factor. The result is best illustrated by the contour plot of the posterior mean where  $X_1$  and  $X_2$  are the two observations and the “estimate” is the posterior mean. (See Appendix 1).

#### 4. THE SIMPLE EXPONENTIAL FAMILY

We now show that the method used to derive simultaneous estimations of the individual premium and of the hyperparameters can be applied to the whole simple exponential family. We use some of the results of JEWELL (1974).



Before defining the new model, we simplify our parameterization. In the basic model, the a posteriori individual premium for risk  $i$ , given the precisions, is

$$E(\mu_i | D, \rho, \pi) = \frac{\pi \cdot m + \rho n_i \bar{x}_i}{\pi + \rho n_i}$$

which can be rewritten in the following way

$$E(\mu_i | D, \rho, \pi) = \frac{\nu m + n_i \bar{x}_i}{\nu + n_i}$$

with

$$\nu = \frac{\pi}{\rho}$$

and it is apparent that only the ratio of  $\pi$  over  $\rho$  is relevant, and not the two variance components separately.

We shall refer to  $\nu$  as to the time constant. It is equal to the number of individual claims records necessary for the credibility factor to be equal to one half.

This new parametrisation is more “natural” than the one introduced in section 2 since in practical applications one often has a priori information on the credibility factor and therefore on the time constant but not necessarily on the two variance components separately. The reason for using both  $\pi$  and  $\rho$  in section 2 is mathematical tractability.

The exponential family likelihoods with the sample mean as sufficient statistics is

$$p(x | \theta) = a(x) c(\theta)^{-1} e^{-\theta x}$$

and their natural conjugate priors are

$$p(\theta) = d(n_0, x_0)^{-1} c(\theta)^{-n_0} e^{-\theta x_0}.$$

JEWELL (1974) shows that under certain regularity conditions

$$E[\mu(\theta)] = \frac{x_0}{n_0} = m$$

The above family is closed under sampling, so that if we observe  $x_{i1}, \dots, x_{in_i}$  for fixed  $\theta_i$  then  $p(\theta_i | D)$  is of the same form with new parameters

$$n'_0 = n_0 + n_i$$

$$x'_0 = x_0 + \sum_j x_{ij}$$

It follows that

$$E[\mu(\theta_i) | D] = \frac{x_0 + \sum_j x_{ij}}{n_0 + n_i} = \frac{n_0 m + n_i \bar{x}_i}{n_0 + n_i}$$

We now generalize the model by introducing a random time constant. The model assumptions are the following:

- 1) The time constant  $\nu$  is a random variable whose distribution function will be specified later.
- 2) Each risk is characterized by an individual parameter  $\theta_i$ .
  - i)  $\theta_1, \theta_2, \dots, \theta_k$  are independent random variables given  $\nu$ ; their common density is  $p(\theta | \nu) = d(\nu, \nu m)^{-1} c(\theta)^{-\nu} e^{-\theta \nu m}$
  - ii)  $(\theta_i, X_{i1}, \dots, X_{in_i})$   $i = 1, \dots, k$  are independent random vectors given  $\nu$ .
- 3) Given  $\theta_i$ ,  $X_{ij}$  ( $j = 1, \dots, n_i$ ) are independent with common density  $p(x | \theta) = a(x) c(\theta)^{-1} e^{-\theta x}$

The posterior mean now is

$$E(\mu(\theta_i) | D) = \int E(\mu(\theta_i) | D, \nu) p(\nu | D) d\nu$$

where

$$E[\mu(\theta_i) | D, \nu] = \frac{\nu m + n_i \bar{x}_i}{\nu + n_i}$$

and therefore

$$\begin{aligned} E[\mu(\theta_i) | D] &= E\left(\frac{n_i}{\nu + n_i} \mid D\right) \bar{x}_i + E\left(\frac{\nu}{\nu + n_i} \mid D\right) m \\ &= z(D) \cdot \bar{x}_i + [1 - z(D)] \cdot m \end{aligned}$$

On the other hand

$$p(\nu | D) = p(D | \nu) p(\nu) \cdot p(x)^{-1}$$

and

$$\begin{aligned} p(D | \nu) &= \int p(D | \nu, \underline{\theta}) p(\underline{\theta} | \nu) d\theta_1 \cdot d\theta_2 \cdot \dots d\theta_k \\ &\propto \prod_{i=1}^k \int c(\theta_i)^{-n_i} e^{-\theta_i \sum x_{ij}} d^{-1}(\nu, \nu m) c(\theta_i)^{-\nu} e^{-\theta_i \nu m} d\theta_i \\ &\propto d^{-k}(\nu, \nu m) \prod_{i=1}^k \int c(\theta_i)^{-(\nu + n_i)} e^{-\theta_i (\nu m + \sum x_{ij})} d\theta_i \\ &\propto \prod_{i=1}^k \frac{d\left(\nu + n_i, \nu m + \sum_j x_{ij}\right)}{d(\nu, \nu m)} \end{aligned}$$

As in our basic model,  $p(D|\nu)$  is in analytic form. This is so because for given  $\nu$   $p(D|\theta, \nu)$  and  $p(\theta|\nu)$  are conjugate priors.

We illustrate our general result with the following.

### Example

We assume a Poisson likelihood and gamma prior.

$$p(x|\theta) = e^{-\theta} \frac{\theta^x}{x!} \quad x = 0, 1, \dots$$

$$p(\theta|\nu) = \frac{\nu^{\nu m}}{\Gamma(\nu m)} \theta^{\nu m - 1} e^{-\theta \nu} \quad \theta > 0$$

and we have

$$p(D|\nu) = \prod_{i,j} \frac{1}{x_{ij}!} \left( \frac{\nu^{(\nu m)}}{\Gamma(\nu m)} \right)^k \prod_{i=1}^k \frac{\Gamma(\nu m + \sum_j x_{ij})}{(\nu + n_i)^{(\nu m + \sum_j x_{ij})}}$$

We have five risks, each with recorded number of claims for two years

$$D = \left\{ \begin{array}{cc} 2 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{array} \right\}$$

using the estimators given in section 1, we have

$$\hat{w} = 0.2$$

$$\hat{b} = 0.325$$

and the time constant is

$$\hat{\nu} = \frac{\hat{w}}{\hat{b}} = 0.615$$

If we compute the likelihood function  $p(d|\nu)$  another picture emerges

$$\hat{\nu}_{MLE} \simeq 9$$

and the likelihood is very flat for  $\nu$  larger than 2, as is apparent from the graph in appendix 2 where we have assumed  $m = 0.6$ . From the shape of the likelihood it is obvious that  $p(\nu|D)$  and therefore  $z(D)$  will heavily depend on the choice of the prior  $p(\nu)$ .

## 5. ARBITRARY MEASURES OF EXPOSURE – THE BÜHLMANN-STRAUB EXAMPLE

Our basic model can be generalized to the case where the measures of exposure are arbitrary. We make the same assumptions as in section 2 except for assumption 3) which now reads: given  $(\mu_i, \rho)$ ,  $X_{ij}$  ( $j = 1, \dots, n_i$ ) are independent random variables with common density  $\varphi(\mu_i, (p_{ij}\rho)^{-1})$ .  $p_{ij}$  are known measures of exposure.

The posterior mean is computed as in section 2.

$$E(\mu_i | D) = \int E(\mu_i | D, \rho, \pi) p(\rho, \pi | D) d\rho d\pi$$

but now we have

$$E(\mu_i | D, \varrho, \pi) = \frac{\pi m + \varrho p_{i.} \bar{x}_{i.}}{\pi + \varrho p_{i.}} = x_i^*$$

with  $p_{i.}$  and  $\bar{x}_{i.}$  as in section 1. The posterior distribution of  $\varrho$  and  $\pi$  is formally as in section 2.

$$p(\rho, \pi | D) \propto p(D | \rho, \pi) p(\rho, \pi)$$

but the likelihood now is

$$p(D | \rho, \pi) \propto \frac{k}{\pi^2} \rho^{\frac{1}{2} \sum_i n_i} \cdot e^{-\frac{1}{2} [\pi \sum (x_i^* - m)^2 + \rho \sum_i \sum_j p_{ij} (x_{ij} - x_i^*)^2]}$$

$$\prod_i (\pi + \rho p_{i.})$$

The proofs are as in section 2. We now reanalyze the data of BÜHLMANN-STRAUB (1970). In their paper the authors give as-if burning costs of seven different excess of loss treaties. For each treaty we have the burning costs from five different years and each treaty is characterized by a measure of exposure and the gross premium income.

The burning costs (in percent of the gross premium income) are as follows:

|        | $i \quad j$ | Year |      |      |      |      |
|--------|-------------|------|------|------|------|------|
|        |             | 1    | 2    | 3    | 4    | 5    |
| Treaty | 1           | 0.0  | 0.0  | 4.2  | 0.0  | 7.7  |
|        | 2           | 11.3 | 25.0 | 18.5 | 14.3 | 30.0 |
|        | 3           | 8.0  | 1.9  | 7.0  | 3.1  | 5.2  |
|        | 4           | 5.4  | 5.9  | 7.1  | 7.2  | 8.3  |
|        | 5           | 9.7  | 8.9  | 6.7  | 10.3 | 11.1 |
|        | 6           | 9.7  | 14.5 | 10.8 | 12.0 | 13.1 |
|        | 7           | 9.0  | 9.6  | 8.7  | 11.7 | 7.0  |

and the gross premium income (in some monetary unit) are

|        |     | Year |    |    |    |     |
|--------|-----|------|----|----|----|-----|
|        | $j$ |      |    |    |    |     |
|        | $i$ | 1    | 2  | 3  | 4  | 5   |
| Treaty | 1   | 5    | 6  | 8  | 10 | 12  |
|        | 2   | 14   | 14 | 13 | 11 | 10  |
|        | 3   | 18   | 20 | 23 | 25 | 27  |
|        | 4   | 20   | 22 | 25 | 29 | 35  |
|        | 5   | 21   | 24 | 28 | 34 | 42  |
|        | 6   | 43   | 47 | 53 | 61 | 70  |
|        | 7   | 70   | 77 | 85 | 92 | 100 |

Bühlmann-Straub compute  $\hat{w}$  and  $\hat{b}$  using the formula given in section 1; they obtain

$$\hat{w} = 209.0 \cdot 10^{-4}$$

$$\hat{b} = 12.1 \cdot 10^{-4}$$

which gives the following estimates for the precisions

$$\hat{\rho} = \hat{w}^{-1} = 48$$

$$\hat{\pi} = \hat{b}^{-1} = 826$$

the time constant being

$$\hat{v} = \frac{\hat{\pi}}{\hat{\rho}} = 17$$

Instead of computing a point estimate we look at  $p(D \mid \pi, \rho)$ . The contour plot of the likelihood is given in appendix 3.

It is seen that the maximum likelihood estimator is approximately

$$\hat{\rho}_{mle} \simeq 50$$

$$\hat{\pi}_{mle} \simeq 500$$

thus giving a much smaller time constant

$$\hat{v} \simeq 10$$

Thus the impact of the variance estimates on the credibility premium of a small treaty can be quite important. A full bayesian analysis would entail specifying a joint prior distribution for  $\pi$  and  $\rho$  and computing the posterior mean through numerical integration. Since this is straightforward it is omitted here.

## REFERENCES

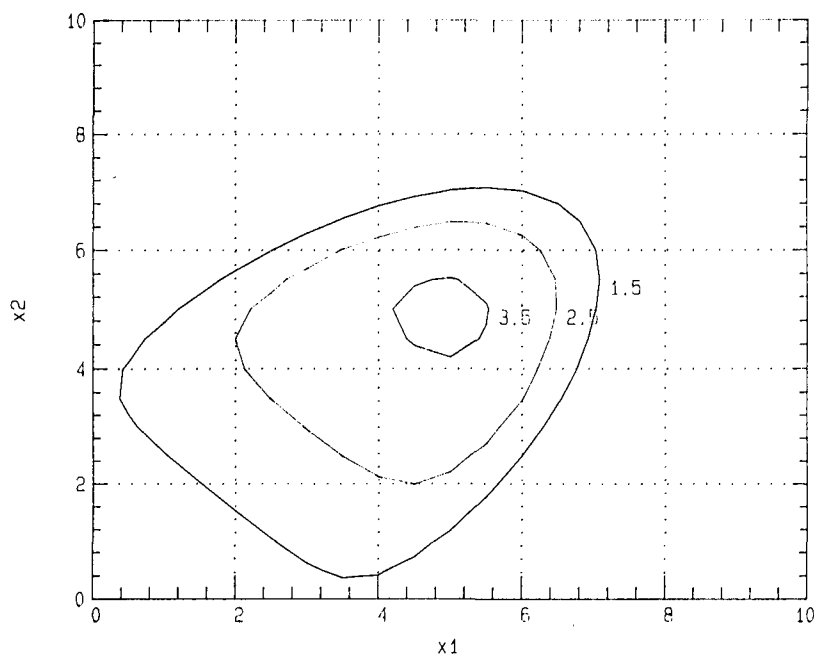
- BÜHLMANN, H. and STRAUB, E. (1970) *Glaubwürdigkeit für Schadensätze*. Mitteilungen der schweizerischen Vereinigung der Versicherungsmathematiker, 1970, pp. 111-113.
- BOX, G. E. P. and TIAO, G. C. (1973) *Bayesian Inference in Statistical Analysis*. Addison-Wesley.
- DUBEY, A. and GISLER, A. (1981) *On Parameter Estimators in Credibility*. Mitteilungen der schweizerischen Vereinigung der Versicherungsmathematiker, pp. 187-212.
- JEWELL, W. S. (1974) *Credible Means are Exact Bayesian for Exponential Families*. ASTIN Bulletin, 8.
- JEWELL, W. S. and SCHNIEPER, R. (1983) *Observation-Dependent Credibility Weights*. Unpublished paper presented at the Berkeley Risk Conference.
- KLUGMAN, S. A. (1986) *Bayesian Credibility with a Noninformative Prior*. Insurance and Risk Theory. Edited by M. Goovaerts, F. de Vylder and J. Haezendonck. Nat. ASI Series C: Mathematical and Physical Science, Vol. 171. Dordrecht, Holland: D. Reidel Publishing Company, pp. 195-206.

RENÉ SCHNIEPER

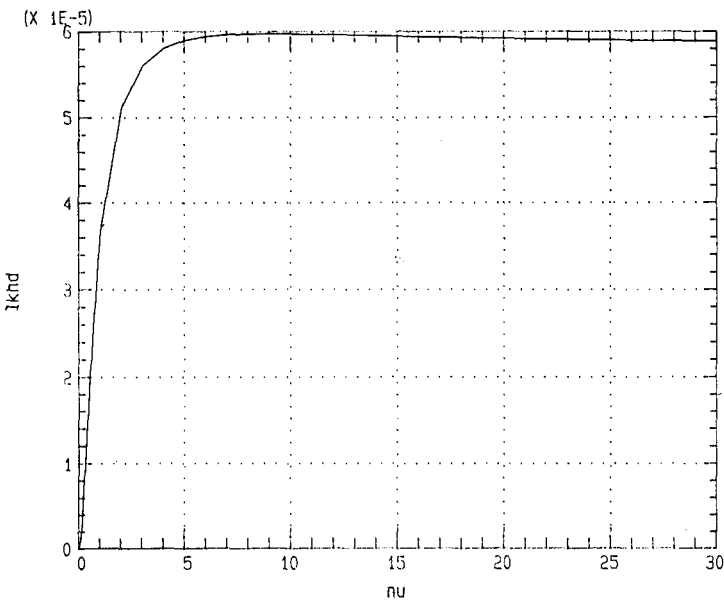
"Zürich" Versicherungs-Gesellschaft

Postfach, CH-8022 Zürich

Appendix 1  
Plot of estimate



Appendix 2  
Plot of  $lkhd$  vs  $\nu$



Appendix 3  
Plot of  $lkhd$

